Teaching Syntax with CLARIN Corpora and Resources

Antonio Balvet
Department of Language Sciences
Univ. Lille, CNRS, UMR 8163 - STL - Savoirs Textes Langage
F-59000 Lille, France
antonio.balvet@univ-lille.fr

Abstract

The recent COVID-19 pandemic has brought online learning to the forefront for learners and teachers. As a consequence, the demand for self-paced and adaptive learning resources has reached unprecedented levels. Fortunately, universities had been using e-learning platforms such as Moodle, or other SCORM-compliant LMS, which has helped make the transition from on-site to on-line learning. However, teachers still have had to design and implement assessment activities in the form of self-correcting activities (true/false, multiple answer questions, mark the words, fill in the blanks questions, etc.). This step has proved to be a major hurdle in the on-site to on-line learning transition, since designing and, most of all, manually editing formative and evaluative assessment activities is a very labour-intensive task. In this article, we present a framework that takes advantage of the corpora and resources available from the LINDAT / CLARIAH-CZ Data & Tools platform in order to generate quizzes and other activities related to syntax. After some background on using NLP for teaching grammar, we present our corpus-to-quiz processing chain, and we outline preliminary results on deploying automatically generated French syntax quizzes in the classroom.

1 Introduction

The domain of education has undergone a profound transformation in recent years, a change dramatically accelerated by the COVID-19 pandemic and the associated lockdown periods. The shift from traditional classroom settings to on-line platforms has not only redefined the way we teach and learn, but also underlined the urgent necessity for deploying innovative educational tools and learning scenarios. In this context, the demand for self-paced, responsive, and personalised learning environments has become more pronounced than ever before.

Even before the pandemic, e-learning platforms such as Moodle were integral to the pedagogical strategies of many universities. Moodle, in particular, allows instructors to include third-party learning applications, thanks to the compliance with inter-operability protocols such as the Sharable Content Object Reference Model (SCORM) or Learning Tools Interactivity (LTI) protocols. These protocols power widgets and other services, such as H5P.org's widgets, or collaborative tools for text or image creation. As a consequence, Moodle is currently being used all around the world, with over 43,000 running instances in Europe alone.[1] Moodle, and other Learning Management Systems (LMS), therefore provided a welcome foundation for the sudden and necessary transition from on-site to on-line learning. However, this transition also exposed significant challenges, particularly in the realm of language education. For instructors, one of the most daunting tasks has been designing and implementing effective assessment activities under strict deadlines. Traditional forms of assessment, such as printed quizzes and tests, have had to be transformed into self-correcting activities such as true/false, multiple choice questions, and other interactive activities. While advantageous in their ability to provide instant feedback and adaptability, such self-correcting activities require a substantial amount of time and effort, especially when catering to large groups of students, who are sometimes facing bandwidth –or even hardware– access restrictions at home. This is particularly true in the context of syntax classes, a fundamental, yet complex, aspect

[1] https://stats.moodle.org/

of linguistics. Syntax poses a unique challenge in an online learning environment, since many students often have unpleasant memories of traditional grammar classes, which focus more on rote memorisation of exceptions rather than understanding underlying principles and parsing methods. In this context, the ability to automatically generate quizzes that are both challenging and adaptive is crucial in a modern educational landscape to enable students to progress at their own pace and according to their individual learning paths. This personalised approach is not just a response to the logistical challenges posed by online education; it is a pedagogical strategy aimed at deepening students' understanding of syntax, enhancing their engagement, and ultimately fostering a more positive and effective learning experience.

In this publication, we outline our proposal to transform the way syntax is taught and evaluated by leveraging the considerable collective research efforts that have gone into Natural Language Processing (NLP) and the extensive manually annotated corpora available from the LINDAT/CLARIAH-CZ Data & Tools platform. One very tangible aspect of our proposal is the design and implementation of a processing chain that converts Universal Dependencies corpora into ready-to-integrate Moodle quizzes. Our processing chain automates the generation of syntax quizzes and other formative activities, letting educators focus on the design of learning scenarios and on providing their students with the necessary feedback, and not on the minutiae of setting up individual quiz questions, thus making the process of deploying formative and evaluative self-correcting activities less prone to error and subjectivity. It also provides a sound foundation for establishing carefully curated collections of standardised syntax tests that educators will be able to share and deploy at their respective teaching sites. Finally, it also offers students a more engaging and stimulating learning environment, thanks to immediate scoring and progression feedback. Although the current version of our corpus-to-quiz processing chain focuses on French syntax quizzes, generated from the French Treebank (Abeillé et al., 2003) and Sequoia (Candito et al., 2014) corpora, indexed on the CLARIN infrastructure and available from universaldependencies.org, the principle presented here is applicable to any Universal Dependencies CONLL-U[2] formatted corpus, with limited overhead. After making the case for syntax "consciousness-raising" and the need for consistent Syntax Competence Levels (SCL) assessment, we provide an outline of projects that have relied on NLP solutions and reference corpora for teaching grammar. We then present our corpus-to-quiz processing chain, and how the generated formative and evaluative activities can be integrated into Moodle courses. Lastly, we report preliminary results of deploying such automatically generated quizzes for both distance and on-site learning, and we offer first insights of deploying standardised automatically generated formative and evaluative activities in the context of our introductory syntax course.

2 Syntax Consciousness-Raising and e-learning

The recent COVID-19 pandemic has highlighted the need for self-paced and adaptive learning resources. Although universities around the world had been using e-learning platforms prior to this event, teachers were still faced with a very labour-intensive task, since designing and implementing self-correcting assessment activities for potentially large groups of learners, in a distance learning context, proved very time-consuming. Furthermore, designing and implementing such assessment activities by hand is error-prone and subjective by nature. Our proposal aims at optimising the design and implementation of such assessment activities by taking advantage of publicly available syntax-annotated corpora. In doing so, we propose a concrete approach to promote syntax 'Consciousness-Raising' among students, thus falling in line with projects such as VISL[3] and the METAL/Gramex presented below.

2.1 From grammar to syntax Consciousness-Raising

The concept of "Consciousness-Raising" (C-R) in grammar teaching, as defined by Rutherford (1987), Schmidt (1990), and Ellis (2016), marked a significant shift in second language acquisition (SLA) pedagogy, especially for English as a Foreign Language (EFL). The status of explicit grammar teaching

[2]CONLL is a tabular format popularised by the Computational Natural Language Learning conferences. CONLL-U corpora are those corpora which follow Universal Dependencies annotation guidelines.

[3]Visual Interactive Syntax Learning project.

has been the subject of intense debates in the last decades, to the extent that many countries have abandoned explicit grammar teaching in SLA and in first language acquisition. However, C-R in grammar can be seen as a transition from a focus on pure grammatical forms to a more nuanced understanding of form-meaning relations: it is often defined as an inductive approach that improves learners' awareness by encouraging them to notice, analyse, and internalise the grammatical aspects of the target language within context, thus fostering a deeper, more meaningful acquisition of language structures. We propose to transpose the notion from the domain of **grammar** to that of **syntax** teaching, in the context of higher education.[4]

We define syntax as the set of rules governing the production and decoding of well-formed sentences in any natural language, associated with an explicit, as well as practical, knowledge of discovery procedures (syntactic tests) and parsing strategies that allow students to look beyond individual surface words and identify the underlying patterns –both at the structural and functional dependency levels– at play in any particular sentence. To this end, we rely on corpus-derived automatically generated exercises, which are described in more detail below. In order to properly achieve syntax C-R, we first need to design and implement sound and consistent learning scenarios designed, precisely, to make students notice, analyse, and internalise syntactic structures, rules and procedures. The proposed learning scenarios, and corresponding sets of self-correcting formative and evaluative syntax activities are designed with a well-defined Syntax Competency Levels (SCL) progression in mind, based on different linguistic features to determine 6 basic SCL, from "beginner" to "intermediate" and "advanced" (see section 2.2).[5] Supplementing traditional "chalk-and-talk" teaching methods with extensive sets of automatically generated quizzes represents a transformative approach to teaching syntax. Our proposal is predicated on the belief that a deep and intuitive understanding of syntax–not just as a set of abstract rules, but as a living structure of communication–is essential for educators who aim to impart linguistic knowledge effectively. For teachers, particularly those instructing in their native language, such a framework offers a dual benefit: it not only reinforces their syntactic proficiency, but also equips them with a methodology to later teach syntax at primary and secondary school levels. By engaging in targeted and adaptive exercises, future teachers can better identify and address their own knowledge gaps, ensuring that they can provide a more comprehensive and nuanced education to their students. For learners of a second language (L2), especially higher education L2 learners, we see targeted exercises focused on syntax competence as an invaluable means of internalising the language they seek to master. Unlike traditional methods that may emphasise the rote memorisation of grammar rules, syntax-focused responsive and adaptive exercises facilitate deeper engagement with the language, encouraging learners to recognise and apply grammatical structures in context.

2.2 Syntax Competence Levels

A critical aspect of our approach is the concept of "Syntax Competence Levels" (SCL). This concept, inspired by the well-established language proficiency levels of the Common European Framework of Reference (CEFR), aims to categorise and define various stages of syntactic understanding and abilities in learners, ranging from A1 (beginner) to C2 (advanced). The British Council's "Grammatical Proficiency Levels" offer a parallel, yet broader, framework, encompassing a wide range of grammatical knowledge including tense, mood, and the usage of prepositions in phrasal verbs. In contrast, "Syntax Competence Levels" concentrate on tasks such as part-of-speech identification, the identification of syntactic structures, and the recognition of functional relationships, preferably in sentences taken from authentic documents rather than forged ones. From a practical standpoint, at the A1 level, the focus is on basic part-of-speech identification and understanding simple sentence structures. This foundational level is crucial for beginners, as it lays the groundwork for more advanced syntactic concepts. As learners progress to higher levels, the complexity of the structures and the depth of syntactic analysis increase.

[4] We address groups of students aiming, in their vast majority, at becoming primary or secondary teachers. These students are native French speakers who often lack a proper understanding of the structural aspects of French syntax.

[5] The definition of SCL is still a work in progress. So far, we have defined 6 SCL, ranging from A1 to C2, based on lexical parameters (e.g. frequency, familiarity), structural (e.g. simple vs complex sentences) and functional features, as well as overall sentence readability and topic.

At the intermediate levels (B1 and B2), learners engage with more complex sentences, delving into intricate syntactic structures and beginning to explore functional relationships between sentence elements (e.g. verb-dependent subordinate clauses, recursive structures). The advanced levels (C1 and C2) challenge learners to master sophisticated aspects of syntax. This includes a deep understanding of nuanced syntactic structures, the ability to analyse and interpret complex functional relationships, and the skill to apply this knowledge in varied linguistic contexts. Such a progression allows for a more targeted approach in teaching and assessing syntax, a fundamental component of linguistic competence that is often challenging for learners. We see this specialisation as particularly beneficial in an academic setting, where a deep understanding of syntax is crucial for students pursuing linguistics or related fields. By implementing these Syntax Competence Levels in our self-correcting activities, we ensure that the generated formative and evaluative activities are not only tailored to the individual learner's current level of understanding, but also provide a clear pathway for progression.

3 Corpora and NLP tools for Computer-Assisted Language Learning

In the domain of Computer-Assisted Language Learning (CALL), several projects have explored the use of reference corpora and NLP tools, to automatically derive vocabulary questions and grammar exercises. These include Fill In the Blank (FIB), Shuffle exercises,[6] true/false and multiple answer question quizzes. Such seminal projects as Bitzer et al. (1961) illustrate how computer technology has, almost from the very start, been seen as a tool to assist in –and optimise– the process of second language acquisition. More recent approaches (Aldabe et al., 2006; Borin & Saxena, 2005; Heck & Meurers, 2022; Lee & Seneff, 2007; Mitkov et al., 2006; Perez-Beltrachini et al., 2012; Smith et al., 2010) are particularly relevant to the present discussion. Following Perez-Beltrachini et al. (2012), we distinguish approaches that rely on reference corpora to extract relevant material from those that rely on some form of word or text generation under specific linguistic constraints.

Our corpus-to-quiz processing chain belongs to the former type of approach. It aims both at reducing manual edition to a minimum and at overcoming the subjectivity (and errors) associated with manually created exercises. Here, we report experiments conducted on French, although the approach can be extended to other languages, since CONLL-U corpora are readily available from universaldependencies.org for over 100 different languages. Our proposal falls in line with a larger ongoing effort to integrate technology –and more recently, Artificial Intelligence– into pedagogy. The integration of technology in syntax C-R has been marked by significant milestones: years before Moodle and other LMS were a widespread commodity, substantial work had been done to leverage Natural Language Processing (NLP) tools for educational purposes. In this context, key projects such as VISL and the LORIA-led Gramex/METAL[7] project are of particular relevance, each contributing uniquely to the advancement of "smart" syntax-aware learning activities.

3.1 The Visual Interactive Syntax Learning project

The VISL[8] project, initiated in the early 2000s, stands as a pioneering effort in this domain. Spearheaded by a consortium led by the University of Southern Denmark, VISL used a Constraint Grammar (CG) parser architecture, designed to accommodate various languages (Bick, 2001, 2004, 2015; Karlsson, 1990; Karlsson et al., 2011; Uibo & Bick, 2005; Wijlff, 2006). Based on this CG parser, large electronic corpora were produced, allowing researchers from the VISL consortium to implement a platform very similar to the well-known "Sketch Engine" presented in Kilgarriff et al. (2008). In addition to syntax-aware concordancers, the VISL consortium also designed several gamified activities based on CG-parsed corpora, for different languages.[9] This project was instrumental in establishing large, syntactically parsed corpora, which formed the basis for a range of syntax-learning tools, among which syntax-aware con-

[6] Fill In the Blank exercises require the user to produce the correct form in a given context, whereas Shuffle exercises require them to produce a valid sentence based on a set of predefined words.

[7] Models and learning analytics for language learning.

[8] Visual Interactive Syntax Learning.

[9] A syntactic labyrinth, as well as a syntactic Tetris and other syntactic games were implemented as Java applets, which means they are unfortunately not functional anymore.

cordancers and an assortment of interactive exercises –some of them gamified– derived from CG-parsed corpora. The innovative approach of the VISL project to grammar learning, emphasising visual and interactive elements, marked a significant departure from traditional syntax teaching methods.

3.2 METAL and Gramex: teaching French grammar with NLP tools

The METAL and Gramex projects, led by LORIA,[10] illustrate another approach to the adoption of NLP tools to teach French grammar. At this point, it is worth underlining that the very first version of Gramex was dedicated to generating "grammar exercises used for language learning i.e., grammar exercises whose syntax and lexicon are strongly controlled" (Perez-Beltrachini et al., 2012). Therefore, it relied on a symbolic, formalism-driven text generation platform to produce sentences based on a set of preestablished patterns to generate FIB and Shuffle grammar exercises. The latest evolution of the Gramex/METAL project saw a change of perspective on the issue of grammar and syntax C-R: the project now relies on high-precision, robust parsers for the automatic parsing of preexisting documents[11] to generate grammar exercises. Notwithstanding these technical changes, the emphasis is still on grammar C-R, with a main target composed of primary and secondary school pupils. This effort contributed significantly to the development of effective automated teaching tools for young learners.[12] In recent years, the Gramex/METAL project expanded its scope by forming collaborations with Metz and Nancy-based primary schools[13] to design comprehensive grammar teaching material associated with targeted exercises. In addition, since the project uses authentic documents (texts and on-line resources), it allows educators to generate exercise questions directly from specific content or URLs, to provide a more engaging and stimulating learning experience. Gramex also addresses collaborative learning by designing exercises that can be carried out by student groups, to foster a more interactive and supportive learning environment. This integrated approach to grammar instruction aims to make grammar learning more accessible, interactive, and effective for primary and secondary school pupils.

4 Corpus-to-quiz: leveraging reference corpora for syntax C-R

The work carried out by VISL and METAL/Gramex represents a significant conceptual foundation on which our project is built. These projects have shown promising perspectives on the use of annotated corpora and NLP for syntax C-R, but unfortunately, the developed applications and activities are no longer functional in the case of VISL,[14] or not yet available in the public domain at the time of writing, in the case of Gramex/METAL. Our approach follows the steps of VISL in that it relies on sentences extracted from existing reference corpora, one important difference being that the corpora used for French exhibit high-quality manually revised syntactic annotations.[15] Our Python-based corpus-to-quiz processing chain, together with a first version of ready-to-integrate part-of-speech and functional relationship identification quizzes, is available from the ACE-Annotated Corpora for online Exercises github repository. Both VISL and Gramex/METAL rely on high-precision parsers, while our approach capitalises on the availability of richly manually annotated corpora from the universaldependencies.org platform. In contrast to Gramex/METAL, our approach aims to address syntax C-R at the university level, targeting a demographic that has largely been overlooked in previous projects: that of aspiring primary and secondary school teachers, as well as French as a Foreign Language instructors. As stated earlier, although French is the main target in its current version, our framework is designed to be adaptable to different languages and educational levels. This adaptability is achieved by leveraging the CONLL-U formatted corpora for Universal Dependencies, making it possible to extend our methodology to a wide range of languages and syntactic structures. Finally, our approach focusses on the exercise generation

[10] Laboratoire Lorrain de Recherche en Informatique et ses Applications.

[11] Authentic documents taken from reference texts, or forged sentences that illustrate a specific grammar topic.

[12] See Perez-Beltrachini et al. (2012) and Colin (2020).

[13] See Gramex.

[14] All developed applications and gamified activities were developed as Java applets, which have been officially deprecated in 2017.

[15] More precisely, the French Treebank has undergone comprehensive manual verification, while a hybrid manual/automatic process has been implemented for Sequoia.

aspect, while all authentication procedures and learning analytics logging are currently being handled by Moodle, thus keeping software infrastructure and personal data management issues to a minimum. Nevertheless, in the coming years, we have plans to develop third-party LTI-compliant web services to specifically overcome limitations in the learning profiling capabilities of most LMS, including Moodle, as well as user interactivity: in its current version, Moodle only natively supports true/false or multiple-answer quizzes, which is a serious limitation for the purpose of our approach to syntax C-R.

4.1 The corpus-to-quiz processing chain

As stated above, our "corpus-to-quiz" processing chain is designed to transform annotated linguistic data into engaging formative and evaluative activities. Our approach relies on CONLL-U formatted corpora, specifically the French Treebank (FTB) and the Sequoia corpora. The French Treebank initiated in the late 1990's, adopting the Penn Treebank methodology.[16] As a consequence, the French Treebank (ca. 700,000 tokens) was essentially a manual annotation project, based on sentences sampled from *Le Monde*, a reference French newspaper. The initial annotations were based on an XML schema in order to capture constituency structure together with functional relationships. It is worth mentioning here that the FTB initiated in a context were dependency annotations were not as widespread as they are now. Moreover, due to copyright restrictions, the corpus is distributed under a specific license.[17] In contrast to the French Treebank, the Sequoia corpus was compiled from different sources: newspaper articles from *L'Est Républicain*, French Europarl samples, French Wikipédia articles, and medication leaflets issued by the European Medicines Agency. Sequoia is, therefore, more varied, and smaller than the FTB, since it totals a little over 70,500 tokens. Both corpora were later processed to accommodate the CONLL-U format and the Universal Dependencies syntactic annotation guidelines.

Figure 1 gives an overview of the main modules of our corpus-to-quiz processing chain.

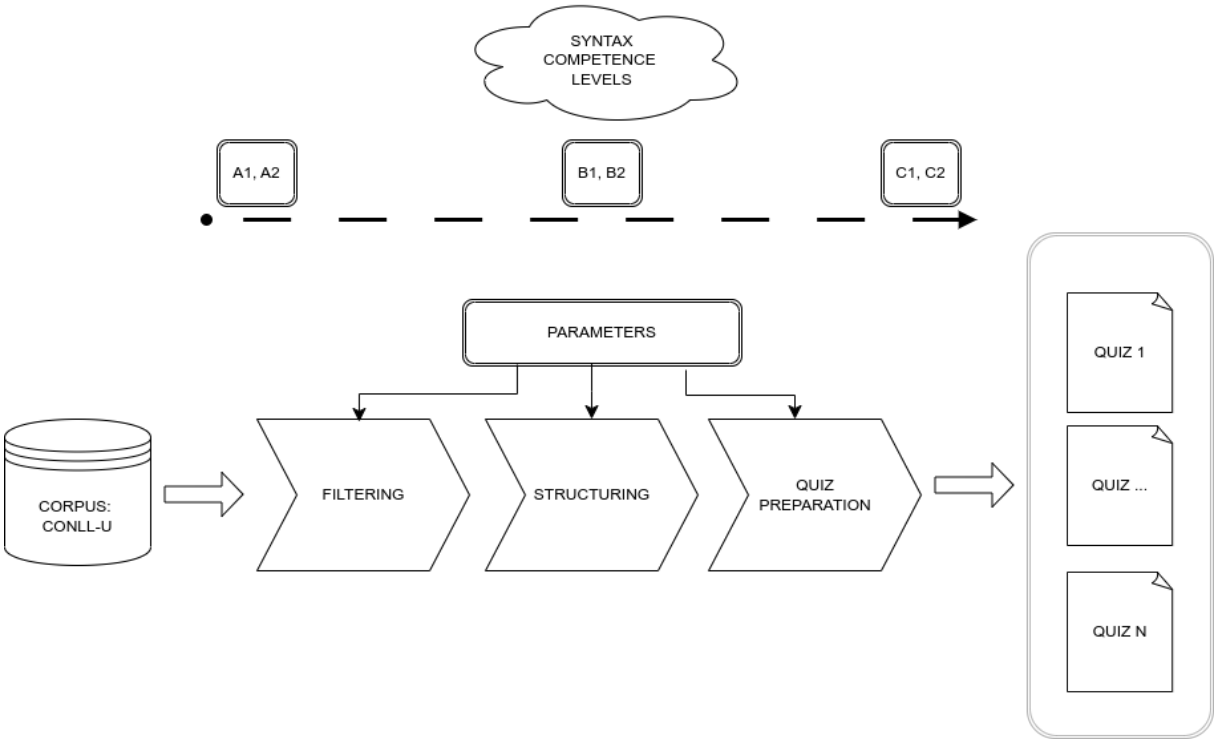

Figure 1: The Corpus-to-Quiz processing chain

As illustrated in Figure 1, the processing chain begins with the selection ('filtering' module) of sentences realising predefined types of syntactic structures from our reference corpora. These sentences are

[16] See Marcus et al. (1993).

[17] As a consequence, the FTB is not directly available from universaldependencies.org, but from the dedicated FTB website, under two distinct licenses (research/commercial).

chosen based on their CONLL-U formatted Universal Dependencies annotations. The underlying Python scripts are the powerhouse of this chain, meticulously processing each sentence to extract its syntactic elements, such as parts of speech, dependency relations, and functional roles. The extracted sentences are then processed to structure and output ready-to-use sets of questions ('structuring' and 'quiz preparation' modules), forming a consistent SCL progression step.

At each step of our processing chain, SCL form the overarching controlling element in that each target syntactic (and morphological) feature can be set at startup, with dedicated execution parameters (see section 4.2). Upon processing, the extracted sentences are then transformed into a structured quiz format, using the General Import Format Template (GIFT), a Moodle-specific text markup formatting. Figure 2 shows an example of an individual question that belongs to a quiz on parts-of-speech.

```
213
214 // question: 8245282  name: Parties du Discours
215 ::Parties du Discours::[markdown]Donner la partie du discours du mot **que** dans la phrase\:\nElle confirme **que**
    l'instauration d'un mécanisme de régulation de l'activité de la médecine libérale n'est pas seulement une Arlésienne.{
216         ~V Subj
217         ~V Part Passé
218         ~PRO Obj
219         ~PRO Int
220         ~V Inf
221         =Conj de Sub
222 }
```

Figure 2: A GIFT-formatted quiz question

As shown in Figure 2, GIFT allows for some level of output formatting (HTML or Markdown) and structuring: the individual question shown is associated with a particular section header ('::Parties du Discours::'), which will ensure that the question will be included in the desired subsection of a Moodle question bank upon importing the GIFT file.[18] Figure 3 shows how this GIFT-formatted question is rendered by Moodle, in the context of a formative or evaluative syntax quiz on parts-of-speech.

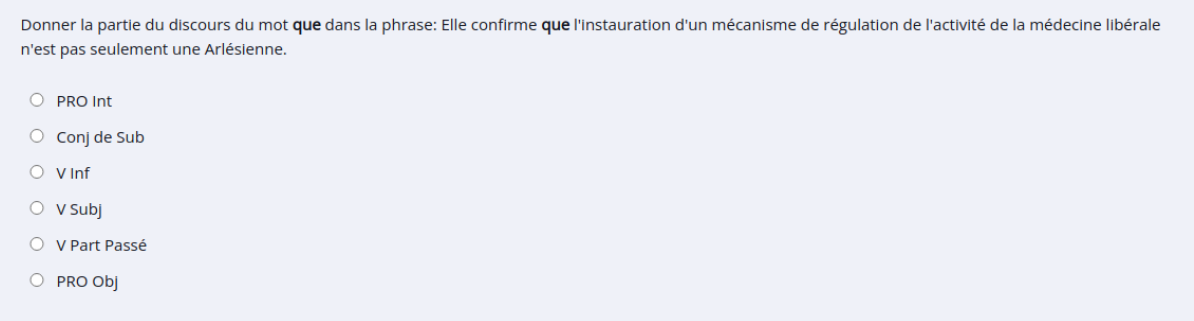

Figure 3: A GIFT-formatted question rendered in Moodle

As can be seen in Figure 3, the set of possible answers is parsed to render the question as a multiple choice quiz with only one correct answer (i.e. 'Conj de Sub' or subordinating conjunction).[19] In this particular case, the set of possible answers is randomly shuffled each time a student accesses the quiz.[20]

The versatility of the GIFT format allows for a wide range of question types — from True/False, to Multiple-Choice and Fill-In-the-Blank questions — making it ideal for testing different aspects of syntactic knowledge with a minimum of formatting and editing. As can be seen, GIFT is also a good trade-off between simplicity and control: it is far more readable than Moodle XML (or other heavily structured formats) while giving quiz managers and final users (i.e. colleagues importing preexisting quiz sets into their own Moodle courses) the necessary control to manage how the questions will be parsed and rendered by Moodle. GIFT-formatted quizzes can easily be imported into Moodle question

[18] Tags can also be associated with individual questions, thus providing a free-text indexing system (e.g. 'simple sentence', 'ditransitive verb', 'readibility:0.8', etc.) of crucial importance when dealing with sets of thousands of individual questions.

[19] An HTML form with radio buttons (only one selection allowed), instead of checkboxes (multiple selections allowed).

[20] Test parameters such as randomised shuffling, total allotted time, number of trials, minimum score to pass, completion status, etc. need to be set manually for each test.

banks, facilitating a streamlined transition from quiz design and generation to deployment in online and in-person classroom settings.[21]

4.2 Controlling the corpus-to-quiz output with SCL-informed parameters

As mentioned above, our scripts use execution parameters provided at startup to control every step of the quiz generation process. These parameters can be used to determine:

- the number of distractors to be included in the generated question
- the particular subset of part-of-speech tags, or of functional relationships
- the way part-of-speech tags and functional relationship labels will be presented to the user
- constraints on the token or on the lemmatized form
- constraints on specific morphological features (tense, mood, number, etc.).

As an illustration, these parameters can be combined to generate a set of individual questions, based on a list of relevant part-of-speech tags such as: prepositions, determiners, adverbs, pronouns (personal and relative) and (coordinating and subordinating) conjunctions, since distinguishing the different grammatical words is often considered a challenging task by A1-A2 SCL students. To make the quiz more engaging, we can not only set the quiz generation parameters to include some nouns, verbs, or adjectives, but also target specific words (or lemmas) belonging to those categories. In the context of syntactic C-R, such features are essential since native French speakers routinely confuse coordinating conjunction “ou” (or) and relative pronoun “où” (where/when), coordinating conjunction “et” (and) and verb “est” (is), or nouns bearing the suffix “-ment” with manner adverbs (e.g. noun “département” and adverb “doucement”).[22] In the domain of syntactic functions, the same principles apply: questions can be generated focussed on specific functional relationships and specific parts of speech. For example, a formative set of quizzes has been set up as preparatory exercises for the end-of-semester final exam. These quizzes were targeted at infinitival subordinate clauses (e.g. “Travailler à l’étranger inquiète mes parents”, working abroad worries my parents), as well as completive subordinate clauses (e.g. “Ils savent qu’un temps d’adaptation sera nécessaire”, they know they will need some time to fit in) that realise either a subject or a direct object function.[23] In the case of these end-of-semester preparatory exercises, participating students had to complete 100 questions in 20 minutes. Each set of 100 questions was randomly sampled from over 2,000 individual questions so that students could extensively test their ability to apply the proposed parsing strategies to new examples taken from a consistent subset of syntactic phenomena. In addition, the quiz was set so that each individual question had a variable number of distractors (between 4 and 8), while a subset of questions targeted at different structures was added to maintain high levels of engagement and attention.[24]

At this point, it is worth underlining that native Universal Dependencies annotations do not fully align with the simplified view of parts-of-speech and functional relationships we use in introductory syntax courses. Therefore, a mapping table has been integrated into our quiz generation scripts: a tag such as “ADP” (adposition) was translated into “Préposition” (preposition), while the function label “nsubj” (nominal subject) was rendered as “Sujet” (subject). These cases are trivial one-to-one correspondences, but since UD annotation conventions follow the ‘universal’ annotation guidelines, meant for comparative syntax studies and not specifically for French syntax, in many cases the transposition requires

[21] Moodle quizzes can be printed on paper for classical pen-and-paper tests, while still retaining the possibility for an (almost) immediate feedback once response sheets have been scanned and processed (typical processing times: 15 minutes for over 200 response sheets, manual corrections included).

[22] “department” and “slowly”. All examples provided are taken from actual syntax quiz questions we routinely use to identify students eligible for syntax tutoring classes.

[23] These cases are particularly challenging for students who have not yet embraced the systematic use of syntactic tests (e.g. pronominalization, passivisation and cleft sentences) and still resort to naïve strategies based on “who did what to whom” questions, in other words, semantic rather than syntactic parsing strategies typical of A1 SCL students.

[24] In our draft SCL progression, such a setting corresponds to A1-A2, which is the SCL required to pass introductory syntax courses, with C1-C2 as the target for completing a full L3 curriculum in linguistics.

some level of shallow parsing of the involved constituents. For example, in French (and many other languages) completive subordinate clauses can realise a verbal root's direct object, e.g.: in "ils savent **qu'un temps d'adaptation sera nécessaire**" the completive subordinate clause can be seen as a direct object of "savent",[25] However, in UD annotations, this constituent will be marked as "ccomp" (clausal complement) and not "obj". Therefore, our mapping table strategy is not just a simple equivalence list, as it must take into account structural constituents as well as functional ones.[26] This level of detail in tailoring syntax quizzes is crucial for promoting a deeper understanding of syntax among students, as it allows instructors to align the generated quizzes with a predefined learning scenario and SCL progression.

5 Discussion and perspectives

5.1 First results and classroom application

We first introduced automatically generated quizzes in our introductory syntax course in 2019. Since then, more than 800 first-year university students have been exposed to these self-correcting exercises. Initially designed as a complement to traditional in-person "chalk-and-talk" teaching methods, these quizzes were originally nothing more than a quick (and admittedly dirty) way to automatically generate more syntax quizzes than could be achieved by hand.[27] However, the unforeseen shift to online learning due to the COVID-19 pandemic propelled these quizzes to the forefront of our syntax education methodology. This transition has not only been a necessity, it has also been an opportunity to rigorously test and refine our approach in a real-world educational context. The COVID-19 pandemic, therefore, represented both a challenge and an opportunity to ramp up the initial concept's development, to arrive at the present situation where we are able to offer a range of curated formative and evaluative activities, providing students with a comprehensive tool for self-assessment and practice throughout the academic year. Figure 4 shows descriptive data on a whole promotion of first year students (n = 170)[28] enrolled in our syntax introductory course (2023-2024 academic year).

The scores for different quizzes are shown: 'POS Quiz' 1 and 2 are formative activities in preparation of mid-term or end-of-semester exams, targeted at parts-of-speech (rated for A1 to B1 SCL). 'TUTOR POS Quiz' is a special offline quiz we use to assess which students seem in need of personalised tutoring, 4 weeks into the semester. As can be seen, performance is rather poor for this particular test, but this can be explained by the fact that most first year students find PoS-tag identification and basic constituent identification tasks challenging at this stage, since these are not routinely taught in secondary school French grammar courses. Finally, 'PREP Func Quiz' is a formative activity on function identification, in preparation of end-of-semester exams (rated for B1 SCL). As is evident from the scores' distributions, a great deal of variation can be observed, which is to be expected if we take into consideration that the different POS quizzes are given at different moments throughout the semester[29] and, as the name suggests, POS Quiz 2 is more stringent than POS Quiz 1. The preparatory quiz on functions is much harder than any of the POS quizzes, since it targets functions such as direct and indirect objects, including cases where subordinate and infinitival clauses realise the targeted functions. In spite of seemingly high student attrition rates,[30] the data show relatively high rates of engagement, since no less than 90 students (53%) engage in any of the proposed quizzes. At this point, it is worth mentioning that those quizzes are not mandatory, even though a small bonus was attributed to the most assiduous students.[31] As a last comment, a clear distinction can be seen between high performers, i.e. students who embrace the proposed methodology, and those who still resort to naïve strategies and thus perform (much) lower.

The results of introducing these new learning activities have been encouraging, so far; students rou-

[25] From a French syntax viewpoint, such a constituent passes the personal pronoun test: "ils **le** savent", which is evocative of a direct object, a valid argument with regard to this verbal root's argument structure.

[26] The underlying strategy is to introduce UD's original annotations and syntactic representation logic at later stages of SCL acquisition.

[27] A first test was made using Perl scripts, which required considerable post-editing of the generated quizzes.

[28] The difference between the total number of enlisted students and the numbers in Figure 4 is due to student attrition.

[29] Every 3 to 4 weeks.

[30] Actually, these figures are 'normal': an average of 30% student attrition is commonly observed every year.

[31] Moodle analytics, based on completion statuses, were essential to assess engagement in this regard.

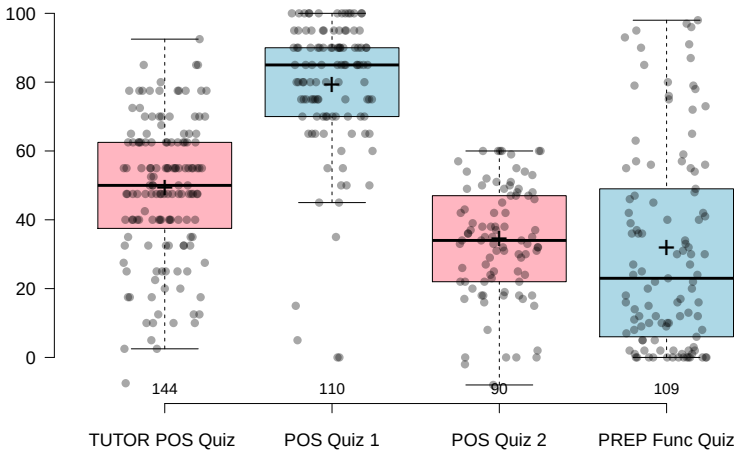

Figure 4: Syntax quizzes scores boxplot (sept. 2023-dec. 2023)

tinely engage with the proposed quizzes, finding them a reassuring tool for preparing for mid-term and end-of-term exams. One of the most important features, in the context of our syntax introductory course, is the instant feedback provided by such self-correcting learning activities. This feature is instrumental in improving student motivation and engagement, a notable improvement over traditional syntax exercises.[32] While there is still considerable room for improvement, especially in catering for less advanced students in their SCL progression, the immediate feedback mechanism not only helps students understand their mistakes in real time, but also allows them to track their progress, fostering a sense of achievement and encouraging continuous learning.

From the instructors' perspective, integrating these quizzes into the curriculum has been equally beneficial. Since our syntax quizzes can be easily imported into existing Moodle courses, they are now being used by other colleagues in charge of other syntax courses, as well as advanced students in charge of one-to-one tutoring classes. Moodle's learner analytics, although not as comprehensive as would be necessary for individual learner profiling, have nevertheless been invaluable in identifying the effectiveness of individual questions and overall quiz structures, allowing us to pinpoint "hard" and "easy" questions and adjust the quiz content accordingly, leading to a more targeted and effective teaching approach. The collected analytics have also provided insights into student learning patterns, aiding in the refinement of the Syntax Competence Levels and the overall curriculum. The quizzes have become more than just an assessment tool; they are now an integral part of our teaching methodology, paving the way for a truly dynamic, interactive, and personalised learning environment.

5.2 Perspectives

In this article, the emphasis has been on generating syntax quizzes from reference corpora. So far, Moodle does not support more interactive learning activities, such as drawing syntactic diagrams, or alternatives to multiple-choice questions (e.g. parsing student submitted hypotheses on the status of a given

[32] Students are receptive to the fact that the proposed quizzes are an experiment, which elicits feedback where, in previous years, none was to be had, positive or otherwise.

constituent).[33] In the context of our project, the emergence of the Learning Tools Interoperability (LTI) protocol[34] has opened new avenues for integrating third-party learning activities, surpassing the capabilities of previous standards such as SCORM (Sharable Content Object Reference Model), supported by well-known platforms such as H5P or Hot Potatoes.[35] LTI stands out as a more versatile and powerful tool, particularly in facilitating the incorporation of third-party web services into learning platforms such as Moodle. The LTI protocol will enable complex processing based on student data within a Moodle (or other LTI-compliant LMS) course. This capability is crucial for our project, since native Moodle activities are limited to variations of Multiple Choice quizzes. Furthermore, although Moodle provides learner activity and performance logging to some extent, no real learner profiling features are available. A fine-grained and comprehensive learner profiling is necessary for the next steps of our project: analysing learner responses and interactions is central to fine-tuning our SCL progression programme. From our standpoint, deploying LTI-compliant third-party smart web services is the way to go to provide a real adaptative, responsive, and personalised learning environment while making sure that the proposed content remains challenging and appropriate for each student's skill level.

Looking toward the future, our plans involve leveraging LTI to its fullest potential. By doing so, we aim not only to enhance the Moodle experience, but also to explore integration with other LMS platforms that support LTI. This broader integration aligns with our goal of making syntax education more accessible and effective across various educational contexts. Moreover, the use of LTI opens up possibilities for developing more sophisticated web services that can perform a range of functions, from learner profiling to complex syntactic analysis. These services can operate as standalone tools or in conjunction with Moodle, providing educators and learners with a versatile and powerful suite of tools for syntax education. In conclusion, the integration of our corpus-to-quiz processing chain into various platforms through the LTI protocol represents a significant step forward in our project. It not only enhances the capabilities of our system within Moodle, it also sets the stage for future expansions and innovations in the field of syntax education technology. By harnessing the power of LTI, we are poised to offer an educational experience that uses new educational tools, with a responsive and personalised learning environment in mind.

References

Abeillé, A., Clément, L., & Toussenel, F. (2003). Building a Treebank for French. *Treebanks: Building and using parsed corpora*, 165–187.

Aldabe, I., De Lacalle, M. L., Maritxalar, M., Martinez, E., & Uria, L. (2006). Arikiturri: an automatic question generator based on corpora and NLP techniques. *Intelligent Tutoring Systems: 8th International Conference, ITS 2006, Jhongli, Taiwan, June 26-30, 2006. Proceedings 8*, 584–594.

Bick, E. (2001). The VISL system: research and applicative aspects of IT-based learning. *Proceedings of the 13th Nordic Conference of Computational Linguistics (NODALIDA 2001)*.

Bick, E. (2004). Grammar for fun: IT-based grammar learning with VISL. *Copenhagen studies in language*, *30*, 49.

Bick, E. (2015). DanProof: pedagogical spell and grammar checking for Danish. *Proceedings of the International Conference Recent Advances in Natural Language Processing*, 55–62.

Bitzer, D., Braunfeld, P., & Lichtenberger, W. (1961). PLATO: an automatic teaching device. *IRE Transactions on Education*, *4*(4), 157–161.

Borin, L., & Saxena, A. (2005). Grammar, Incorporated. In P. J. Henrichsen (Ed.), *CALL for the nordic languages* (pp. 125–145). Samfundslitteratur. https://samfundslitteratur.dk/bog/call-nordic-languages

[33] Keeping cheating in check is not the only concern. Students must be able to explain how and why they arrive at a given response.

[34] Released in its first version in 2008, by IMS Global, a not-for-profit learning consortium.

[35] Launched in 1995, the SCORM standard is developed by the Advanced Distributed Learning (ADL) consortium, with main contributors from the IMS (Instructional Management System) project.

Candito, M., Perrier, G., Guillaume, B., Ribeyre, C., Fort, K., Seddah, D., & de La Clergerie, É. V. (2014). Deep syntax annotation of the Sequoia French Treebank. *International Conference on Language Resources and Evaluation (LREC)*.

Colin, É. (2020). Traitement Automatique des Langues et génération automatique d'exercices de grammaire. *Theses, Université de Lorraine*.

Ellis, R. (2016). Grammar teaching as Consciousness Raising. *Teaching English grammar to speakers of other languages*, *2*(1), 128–150.

Heck, T., & Meurers, D. (2022). Parametrizable exercise generation from authentic texts: effectively targeting the language means on the curriculum. *Proceedings of the 17th Workshop on Innovative Use of NLP for Building Educational Applications (BEA 2022)*, 154–166.

Karlsson, F. (1990). Constraint Grammar as a framework for parsing running text. *COLING 1990 Volume 3: Papers presented to the 13th International Conference on Computational Linguistics*.

Karlsson, F., Voutilainen, A., Heikkilae, J., & Anttila, A. (2011). *Constraint Grammar: a language-independent system for parsing unrestricted text* (Vol. 4). Walter de Gruyter.

Kilgarriff, A., Rychly, P., Smrz, P., & Tugwell, D. (2008). The Sketch Engine. *Practical Lexicography: a reader*, 297–306.

Lee, J., & Seneff, S. (2007). Automatic generation of CLOZE items for prepositions. *Eighth Annual Conference of the International Speech Communication Association*.

Marcus, M., Santorini, B., & Marcinkiewicz, M. A. (1993). Building a large annotated corpus of English: the Penn Treebank. *Computational linguistics*, *19*(2), 313–330.

Mitkov, R., Le An, H., & Karamanis, N. (2006). A computer-aided environment for generating multiple-choice test items. *Natural language engineering*, *12*(2), 177–194.

Perez-Beltrachini, L., Gardent, C., & Kruszewski, G. (2012). Generating grammar exercises. *Proceedings of the Seventh Workshop on Building Educational Applications Using NLP*, 147–156.

Rutherford, W. E. (1987). The meaning of grammatical Consciousness-Raising. *World Englishes*, *6*(3), 209–216.

Schmidt, R. W. (1990). The role of consciousness in Second Language Learning1. *Applied linguistics*, *11*(2), 129–158.

Smith, S., Avinesh, P., & Kilgarriff, A. (2010). Gap-fill tests for language learners: corpus-driven item generation. *Proceedings of ICON-2010: 8th International Conference on Natural Language Processing*, 1–6.

Uibo, H., & Bick, E. (2005). Treebank-based research and e-learning of Estonian syntax. *Proceedings of Second Baltic Conference on Human Language Technologies: Second Baltic Conference on Human Language Technologies*, 4–5.

Wijlff, A. (2006). VISL in Danish schools. *English Teaching: Practice & Critique (University of Waikato)*, *5*(1).